

AD-A194 333

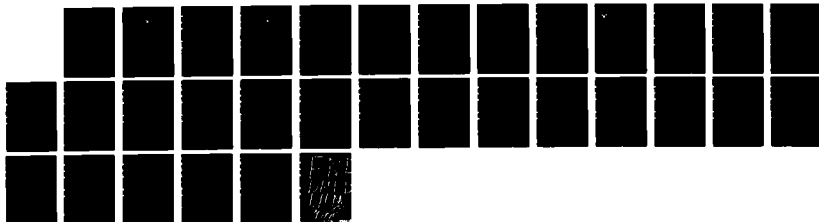
IMPROVING SAC'S INSPECTOR GENERAL AIRCREW TESTING  
PROGRAM(U) AIR COMMAND AND STAFF COLL MAXWELL AFB AL  
E T DIXON APR 88 ACSC-88-0755

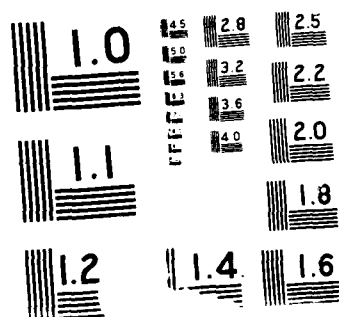
1/1

UNCLASSIFIED

F/G 5/9

NL





AD-A194 333



DTIC  
ELECTE  
JUN 09 1988  
S H D

# AIR COMMAND AND STAFF COLLEGE

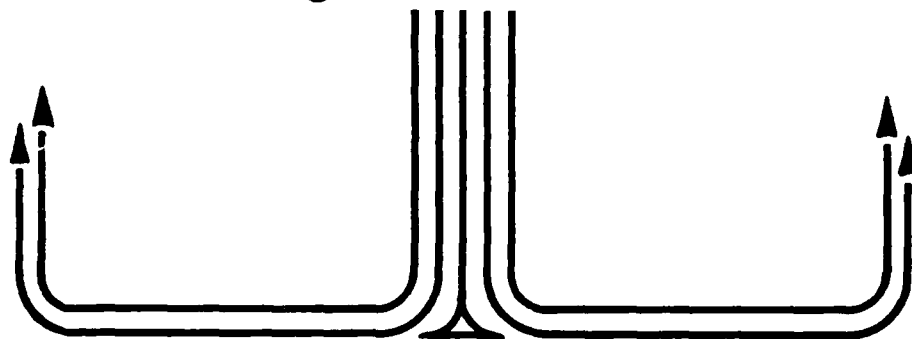
## STUDENT REPORT

IMPROVING SAC'S INSPECTOR GENERAL AIRCREW  
TESTING PROGRAM

MAJOR EDWARD T. DIXON

88-0755

"insights into tomorrow"



# DISCLAIMER

The views and conclusions expressed in this document are those of the author. They are not intended and should not be thought to represent official ideas, attitudes, or policies of any agency of the United States Government. The author has not had special access to official information or ideas and has employed only open-source material available to any writer on this subject.

This document is the property of the United States Government. It is available for distribution to the general public. A loan copy of the document may be obtained from the Air University Interlibrary Loan Service (AUL/LDEX, Maxwell AFB, Alabama, 36112-5564) or the Defense Technical Information Center. Request must include the author's name and complete title of the study.

This document may be reproduced for use in other research reports or educational pursuits contingent upon the following stipulations:

- Reproduction rights do not extend to any copyrighted material that may be contained in the research report.

- All reproduced copies must contain the following credit line: "Reprinted by permission of the Air Command and Staff College."

- All reproduced copies must contain the name(s) of the report's author(s).

- If format modification is necessary to better serve the user's needs, adjustments may be made to this report--this authorization does not extend to copyrighted information or material. The following statement must accompany the modified document: "Adapted from Air Command and Staff College Research Report \_\_\_\_\_ (number) \_\_\_\_\_ entitled \_\_\_\_\_ (title) \_\_\_\_\_ by \_\_\_\_\_ (author)."

- This notice must be included with any reproduced or adapted portions of this document.



**REPORT NUMBER** 88-0755

**TITLE** IMPROVING SAC'S INSPECTOR GENERAL AIRCREW TESTING PROGRAM

**AUTHOR(S)** MAJOR EDWARD T. DIXON, USAF

**FACULTY ADVISOR** MAJOR STEVEN L. HANSEN, ACSC/3824 STUS

**SPONSOR** MAJOR MIKE GENTRY, SAC/DOSK

Submitted to the faculty in partial fulfillment of  
requirements for graduation.

**AIR COMMAND AND STAFF COLLEGE  
AIR UNIVERSITY  
MAXWELL AFB, AL 36112-5542**

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

## REPORT DOCUMENTATION PAGE

Form Approved  
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION / AVAILABILITY OF REPORT STATEMENT "A" Approved for public release; Distribution is unlimited.		
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE					
4. PERFORMING ORGANIZATION REPORT NUMBER(S) 88-0755			5. MONITORING ORGANIZATION REPORT NUMBER(S)		
6a. NAME OF PERFORMING ORGANIZATION ACSC/EDC		6b. OFFICE SYMBOL (if applicable)	7a. NAME OF MONITORING ORGANIZATION		
6c. ADDRESS (City, State, and ZIP Code) MAXWELL AFB AL 36112-5542			7b. ADDRESS (City, State, and ZIP Code)		
8a. NAME OF FUNDING / SPONSORING ORGANIZATION		8b. OFFICE SYMBOL (if applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER		
8c. ADDRESS (City, State, and ZIP Code)			10. SOURCE OF FUNDING NUMBERS		
			PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.
			WORK UNIT ACCESSION NO.		
11. TITLE (Include Security Classification) IMPROVING SAC'S INSPECTOR GENERAL AIRCREW TESTING PROGRAM					
12. PERSONAL AUTHOR(S) Dixon, Edward T., USAF					
13a. TYPE OF REPORT		13b. TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Year, Month, Day) 88 April	15. PAGE COUNT 30
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP			
19. ABSTRACT (Continue on reverse if necessary and identify by block number) SAC has restructured its tactics program and the IG testing process has not kept pace. This study investigates ways for the command to improve this testing program to more accurately measure aircrew knowledge. The author concludes the command can implement several innovations to improve the testing system with little or no expense.					
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a. NAME OF RESPONSIBLE INDIVIDUAL ACSC/EDC Maxwell AFB AL 36112-5542			22b. TELEPHONE (Include Area Code) 293-2867		22c. OFFICE SYMBOL

## PREFACE

The Strategic Air Command (SAC) is in the middle of the largest conceptual change in the command's history. Beginning in 1985, the command has restructured the way it thinks about tactical matters. Driven by the increasing Soviet threat, SAC rewrote its tactical doctrine and placed greater demands upon its aircrews to learn new information. The Inspector General's (IG) testing system is a blanket approach, not reflecting sound educational testing guidance. Tests remain basically the same in form and content as before the new tactics initiatives. The author's purpose is to examine ways for the IG to improve the testing process, accurately measure aircrew knowledge, and assess whether the crews are ready to perform their mission. The author will examine testing history, discuss the attributes of a good examination, examine how other commands and services test, discuss non-written evaluation systems, and, finally, make recommendations.

During this project Major Mike Gentry's assistance from SAC Headquarters was instrumental in determining command testing policy, implementation feasibility, and insights of various general officers. The guidance, patience, and plain honesty of Major Steve Hansen, my advisor, is what kept this paper on time and on target.



Accession For

NTIS GPA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	

By \_\_\_\_\_

Dissemination/

Availability Codes

Dist. \_\_\_\_\_

10-1

## ABOUT THE AUTHOR

### MAJOR EDWARD T. DIXON

Major Dixon received his Bachelor of Arts Degree in Economics from the University of Connecticut in 1976. He was commissioned thorough the AFROTC program after being named a Distinguished Graduate. Major Dixon is a dual rated officer, completing navigator training in 1977 at Mather AFB, California, and pilot training in 1982 at Columbus AFB, Mississippi.

Major Dixon's operational experience is extensive. Upon graduation from navigator training and completion of combat crew training he was assigned to FB-111 bombers at Pease AFB, New Hampshire. Upgrading to instructor radar navigator while still a first lieutenant, he became the unit's expert on tactics and the new electronic countermeasures equipment being installed in the FB-111. He participated in two "Tiger Meet" tactical flying competitions in NATO before his selection to attend pilot training.

In August of 1982, he was an Outstanding Graduate of B-52 combat crew training at Castle AFB, California. Assigned to the 337th Bombardment Squadron at Dyess AFB, Texas, he served as a B-52H copilot and the unit's flying safety officer. He upgraded to aircraft commander in the minimum time allowed and was reassigned to Andersen AFB, Guam to fly B-52Gs. His additional duties at Guam included: self inspection monitor, OER monitor and chief of the instrument flying school. Major Dixon upgraded to instructor pilot in minimum time, returning to Guam to assist in developing B-52 delivery tactics for the AGM-84 Harpoon anti-ship missile. The final position he held before his assignment to ACSC was as the squadron's chief instructor pilot.

Major Dixon has over 3100 hours of flying time including 500 hours in the FB-111, over 2000 hours in the B-52 and 300 hours in the T-37. He has ten years of SAC experience and participated in more than 15 operational readiness inspections. He is a graduate of SOS and ACSC by correspondence and has an MBA in Management from Golden Gate University.



# TABLE OF CONTENTS

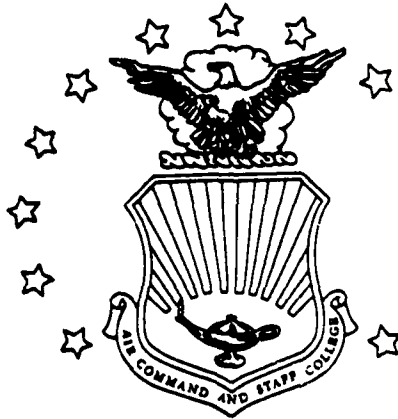
Preface.....	iii
About the Author.....	iv
List of Illustrations.....	vi
Executive Summary.....	vii
CHAPTER ONE--BACKGROUND FOR CHANGE	
Introduction.....	1
Definitions.....	1
Reasons for Change.....	1
Overview.....	2
CHAPTER TWO--A HISTORY OF EVALUATION	
Written Evaluations.....	3
SACR 3-1 Testing Changes.....	3
Inflight Evaluation.....	4
Testing Limitations.....	4
Summary.....	5
CHAPTER THREE--TECHNICALITIES OF TEST CONSTRUCTION	
Learning.....	6
Reasons for Evaluation.....	6
Writing a Test.....	7
Types of Questions.....	7
Summary.....	9
CHAPTER FOUR--HOW OTHERS EVALUATE	
US Army.....	10
US Navy.....	10
MAC.....	10
TAC/USAFE/PACAF.....	11
Summary.....	11
CHAPTER FIVE--NON-WRITTEN EVALUATION	
Formal Oral Evaluations.....	12
Mission Planning.....	13
Using the Simulator.....	13
Informal Oral Evaluations.....	13
Inflight Evaluation.....	14
Summary.....	15
CHAPTER SIX-- SOLUTIONS AND RECOMMENDATIONS	
Tests.....	16
Non-Written Evaluation.....	17
Inflight Evaluation.....	17
Recommendations.....	18
BIBLIOGRAPHY.....	19

---

## LIST OF ILLUSTRATIONS

---

Table 1--Cognitive Domain.....	7
--------------------------------	---



## EXECUTIVE SUMMARY

Part of our College mission is distribution of the students' problem solving products to DoD sponsors and other interested agencies to enhance insight into contemporary, defense related issues. While the College has accepted this product as meeting academic requirements for graduation, the views and opinions expressed or implied are solely those of the author and should not be construed as carrying official sanction.

*"insights into tomorrow"*

**REPORT NUMBER** 88-0755

**AUTHOR(S)** MAJOR EDWARD T. DIXON, USAF

**TITLE** IMPROVING SAC'S INSPECTOR GENERAL AIRCREW TESTING PROGRAM

**Purpose:** To propose improvements to Strategic Air Command's (SAC's) Inspector General aircrew testing program in order to more accurately measure aircrew tactical knowledge, and the aircrew's capability to carry out wartime missions.

**Problem:** SAC is in the midst of a major tactics revision emphasizing threat knowledge and flying skills. New threat systems and improved enemy capabilities are expanding aircrew knowledge requirements. The IG's aircrew evaluation system has not kept pace with the tactics revision program, and needs to be improved to accurately measure expanding aircrew knowledge.

**Discussion:** While he was Commander-in-Chief of SAC, General Larry Welch directed a command wide tactics improvement program. Facing an improved Soviet fighter threat with outdated B-52 electronic countermeasures and a crew force that was rapidly losing experience through retirement, B-1 manning, upgrading pilots and navigators early and airline hiring, General Welch felt that current tactics were merely procedural.

## CONTINUED

The Operational Readiness Inspection (ORI) evaluates SAC aircrew's ability to perform their wartime tasks. The ORI does not yet take into account the new tactics training program. With one exception, the entire ground testing scheme remains unchanged. Test construction guidance in AFM 50-62, Handbook for Air Force Instructors, is not used. Tests are written using guidance found in SAC regulations supporting the aircrew standardization program. This restrictive SAC guidance eliminates many types of questions applicable to ORI testing and builds guessing bias into the test. The inflight portion of the ORI does not evaluate any new flying training events created through the tactics training program.

Our sister services and other MAJCOMs have effective ways to evaluate their aircrews. Aircrews plan missions, drop cargo or weapons, are questioned by evaluators, and take a battery of tests including aircraft recognition. No one organization uses the simulator in evaluation.

Written tests need to be complemented with other evaluation tools. Aircrews briefing specific wartime sorties to an IG panel perform at high cognitive levels. Planning a strike mission uses tactical skills at the same high cognitive levels.

Using the B-52 simulator is not practical because some units have to commute to one, IG evaluators may not be B-52 qualified, and the simulator is used in evaluating tactical knowledge during a different evaluation. Having an IG representative informally quiz aircrews repeats what aircrews have demonstrated in formal briefings while introducing a possible grading bias. These two evaluation means represent redundant evaluation.

Conclusions: The new evaluation program should test tactics separately. A separate grading area for tactical knowledge should be created. The tests given to aircrews should contain many selection type questions and include aircraft recognition. Tests should be written by the 436th Strategic Training Squadron (STS), the unit responsible for writing all tactics training lessons. Tactics questions would be entered into a master question data base for ease in updating information or generating tests via microcomputer.

Non-written evaluation will consist of two formal mission briefings forming 20 percent of the tactical knowledge grade. One aircrew will plan a strike mission to be evaluated tactically. This mission plan will be worth 10 percent of the tactical knowledge grade.

The IG needs to evaluate 50 percent of the aircraft a wing flies during an ORI. This provides a comprehensive look at the aircrews while providing accurate input data for the unit's overall IG rating. Increasing inflight evaluated activities challenges aircrews while providing a more

## CONTINUED

realistic combat simulation. Reducing radio and electronic emissions, flying at night and evaluating weapon release parameters can be incorporated into the evaluation immediately.

Recommendations: SAC's IG testing program can be improved by improving the quality of the evaluations, grading tactics as a separate area, and complimenting the test with formal briefings and a mission planning exercise. Increasing the number of evaluators inflight will enable SAC to gather more accurate data to compute the final inspection rating. Increasing the number of inflight areas evaluated will simulate a realistic combat environment and provide a better picture of the aircrew's capabilities. The author recommends that SAC adopt these proposals.

## Chapter One

### BACKGROUND FOR CHANGE

#### INTRODUCTION

The ability to apply tactical knowledge automatically in an unforgiving threat environment is the key to survival for Strategic Air Command's (SAC) aircrews in the 1990s. Expanding worldwide US commitments, an increasing emphasis on contingency operations, and innovative new missions for SAC's B-52 force provide new challenges for aircrews.

The author's purpose is to investigate ways of enhancing SAC's Inspector General (IG) aircrew testing process to more accurately measure B-52 aircrew tactical knowledge.

#### DEFINITIONS

Webster defines tactics in two ways: "The art or skill of employing available means to accomplish an end", or "The science and art of disposing and maneuvering forces in combat" (3:1186). JCS Publication 1 likewise defines it in two ways: "The employment of units in combat", or "The ordered arrangement and maneuver of units in relation to each other and/or to the enemy in order to utilize their full potentialities" (4:363). The author uses these definitions to shape his own definition of tactical knowledge as knowledge maximizing your capabilities while minimizing those of your opponent.

The author defines the evaluative process as any series of tests, written, oral, or demonstrative, administered on the ground or in flight, designed to measure tactical knowledge.

Emergency War Order (EWO) knowledge is any knowledge, other than tactical knowledge, a B-52 crewmember needs to carry out an assigned mission. An example of EWO knowledge would be safe passage procedures.

#### REASONS FOR CHANGE

Air Force Chief of Staff, General Larry Welch, when serving as Commander in Chief of SAC, recognized the need for tactics reform by calling the current tactics manuals "procedural", and told SAC's Director of Operations, "[We] need to create much better understanding of threat capabilities and how to counter [them]" (21:--).

General Welch was also concerned about the increasing Soviet threat. During the early 1980s the Soviets fielded three new interceptors with a true look down, shoot down capability. In the 1985-86 edition of Jane's All the World's Aircraft, Donald Latham, Assistant Secretary of Defense for C3I, said of the MIG 31: " [It] has better avionics, a better C3 system to work into, a better air to air missile, is faster, has greater combat range and [the Soviets] are producing it like gangbusters" (2:53). Aircrews planning to penetrate Soviet airspace were faced with greater peril than at any other time in SAC's history.

The bomber crews found themselves on the wrong side of the technological cycle. When a major new threat like the MIG 31 is introduced, the defense industry cannot respond and upgrade electronic warfare equipment quickly. Soviet systems need to be analyzed, contracts let, and defensive systems designed, tested, and installed. Until new equipment becomes available, increasing aircrew knowledge and developing new tactics are the only means at SAC's disposal to attempt to counter the Soviet threat.

By the early 1980s, bomber tactics were outdated. Chapter 7 of SACR 55-21, Volume II, B-52 Aircrew Tactical Doctrine, contained little threat specific guidance (9:Ch 7). The threat guidance provided tried to categorize every situation into a universal set of rules to use in defeating a family of threats (9:Ch 7). In the five years the author worked with SACR 55-21, Volume II, there were no major tactics revisions, even though the Soviets were improving their defenses.

Each year the B-52 force grows younger through the loss of experience. Navigators upgrade to radar navigator in 400 hours or less (8:Flg 2-1). The B-1B takes only highly experienced B-52 crewmembers for its crew force (13:4-6). The Advanced Technology Bomber will most likely require the same experience levels. Deregulation brought a hiring boom in the airline industry and some experienced pilots left to start a new life. The core of the Vietnam experience is retiring, leaving SAC with few beneath the rank of lieutenant colonel having actually fired a shot in combat.

## OVERVIEW

This chapter defined tactics, evaluation systems, and discussed why SAC initiated changes to its tactics program. Chapter Two examines the history of the IG's testing program and the tools used to evaluate aircrews. Chapter Three discusses how evaluators effectively measure learning by discussing knowledge levels, cognitive theory, and test design. Chapter Four examines testing programs used by other services and MAJCOMs. Chapter Five examines non-written evaluation tools, such as oral evaluations, the use of the simulator, mission planning exercises, and increasing inflight evaluation. Chapter Six proposes the author's recommendations for improvement.

## Chapter Two

### A HISTORY OF EVALUATION

#### WRITTEN EVALUATIONS

SAC IG written test formats are the same now as when this author entered the command in 1977. Prior to the introduction of SACR 3-1, B-52 Aircraft Tactics, aircrews took a situational tape test to measure their knowledge of JCS execution procedures. This was followed by an eight to ten question test of memory items from SACR 55-21, Volume II. The examination concluded with a 25 question open-book examination covering all regulations and publications carried inflight.

The Command and Control Procedures testing remains an important part of the aircrew testing program since it deals with nuclear weapons. This examination is required by SACR 55-45, Volume V, as a portion of the wing's nuclear surety evaluation (10:9-1). Allowing individuals to take the examination as a crew approximates the mission environment. Every bomber crewmember must be intimately familiar with command and control procedures. It is critical in nuclear surety and must remain in any future testing plan.

Closed book testing was limited to critical memory items in SACR 55-21, Volume II. The test guidance for this examination was traditionally interpreted by aircrews to require a verbatim response. This test measured an individual's ability to memorize large amounts of material instead of applying knowledge to a specific situation or carrying an answer to a logical conclusion.

The open-book test required short answer responses that were generally written verbatim from the source. Answers did not apply knowledge or require a developed, logical conclusion. Instead, aircrews wrote down the response directly from the source regulation. This test measured the aircrew's ability to find 25 answers in 90 minutes rather than measuring how aircrews applied tactical knowledge. The open-book test did contain some demonstration testing by having pilots compute aircraft performance problems and navigators compute force timing and deconfliction problems.

#### SACR 3-1 TESTING CHANGES

Publication of SAC Regulation 3-1 changed testing only slightly. For three months there was a moratorium on IG testing to allow aircrews to familiarize themselves with the new regulation (19:1).



During the next 12 months testing remained unchanged (19:1-2), allowing SAC to assess question validity and the wings to train the aircrews. After 12 months, the entire SACR 3-1 became testable on a closed-book basis (19:2). The only effective change made was including open-book tactics questions in the closed-book section of the examination without changing any test format.

### INFLIGHT EVALUATION

Inflight evaluation evolves continuously. During the 1970s, few flying areas were evaluated. The wing's rating was based on the proportion of accurate weapons delivered versus total weapons. No actual weapons were ever dropped. Scoring was done by radar scoring the aircraft position in relation to the target and applying weapon ballistics to the aircraft position.

The command's damage expectancy (DE) criteria, today's evaluation standard, was introduced in the late 1970s and refined in the early 1980s. DE tries to measure random factors impacting aircrews on a combat mission, such as time control, equipment performance, and navigation accuracy. DE criteria also significantly reduced the allowable weapon miss distance. A few IG observers fly to record the data used to compute the DE rating. Most DE data comes from wing operations and maintenance reports.

### TESTING LIMITATIONS

The ORI scenario allows for up to three aircrew testing sessions. During two ORIs the author witnessed, the IG brought only one test. The first crews took the test and debriefed the squadron commander. An aircraft commander meeting followed this debriefing, compromising the test. The IG rating for those years reflected the squadron's information distribution effectiveness rather than aircrew knowledge.

SAC considers the ORI flight to be just another training mission. Peacetime flying and safety rules apply. There is no enemy trying to kill you. Weapons delivery parameters are not critical and electronic countermeasures maneuvering is not accomplished to insure an accurately scored aircraft position for bomb scores. The author believes flying the ORI mission is not like flying a combat mission.

The ORI flying mission is evaluated according to a set of criteria. Not leaving the low level corridor, receiving 10,000 pounds of fuel from the KC-135 tanker, and maintaining mission timing within two minutes are examples of criteria objectives. If SAC is to keep the flying phase criterion-related and not result oriented, the command should evaluate more criteria to try to more closely measure the skills required for combat.

### SUMMARY

The IG's evaluation program has slowly changed. Simple examinations

still require rote answers, are not carried to logical conclusions, and do not relate to the objectives of the individual tactics lessons. The changes to the testing program, associated with publishing SACR 3-1, simply moved one question category leaving the testing format unchanged. Poor evaluation scheduling, with several testing sessions, can enable aircrews to compromise the test. Few inflight areas are evaluated today and then only if a crew obtains or fails to obtain a very specific criterion.

## Chapter Three

### TECHNICALITIES OF TEST CONSTRUCTION

#### LEARNING

Webster defines learning as the ability "to acquire a skill, knowledge or a behavioral tendency" (3:654). The Air Force student-centered instructional system reflects this definition. Student-centered instruction "describe(s) learning in terms of outcomes rather than instructor activity" (5:3-1). If learning has been accomplished, the student will be able to demonstrate knowledge through a test.

The Air Force feels "learning is best explained by one or some combination of two theories, behaviorism or cognitive theory" (5:2-1). Behaviorists believe that people learn by having behavior reinforced (5:2-1), while cognitive theorists see motivation, generalizing, insight, and discovery as significant concepts (5:2-1). Cognitive theory requires students to go beyond simple recall and gain an understanding of the subject (5:2-1). Dr. Benjamin Bloom's cognitive domain relates given learning levels with appropriate mental activities. The Air Force uses Dr. Bloom's domain to develop learning objectives.

Bloom's taxonomy is reproduced in Table 1. Evaluation represents the most complex cognitive behavior. Learning and mental activity become less intense down to knowledge, the least complex level. Instructors use cognitive domain to specify behaviors (objectives) the student must demonstrate before an evaluator can accept learning has occurred.

#### REASONS FOR EVALUATION

In the Air Force, "our primary reason for evaluation is usually to determine whether students have achieved our stated objectives" (5:20-5). Air Force tests perform other functions beside measuring learning. Staff officers use test results to validate the basic assumptions the tactics training program is based on, or provide feedback to revise the tactics curriculum (1:3). The IG can use test results to evaluate the efficiency of a wing's instructor force or stimulate aircrews to learn new material (1:6-7).

Because IG test results are important for other inputs to the tactics program besides a wing's IG rating, the author feels SAC needs to construct a technically accurate explanation, using the guidance in AFM 50-62, to guarantee accurate results.

<u>LEVEL OF LEARNING</u>	<u>MENTAL ACTIVITY</u>
Evaluation	Exercise learned Judgement
Synthesis	Create new relationships
Analysis	Determine relationships
Application	Use generalizations in specific instances
Comprehension	Translate, Interpret
Knowledge	Recall and recognition

Table 1. Cognitive Domain

#### WRITING A TEST

AFM 50-62 describes the characteristics making up a good test. A technically correct test takes into account reliability, validity, objectivity, comprehensiveness, and must be capable of differentiating (5:20-2). Reliability means the test yields consistent results each time it is administered (5:20-2). A valid test measures exactly what it is supposed to measure (5:20-3). An objective test is able to eliminate the bias of the grader (5:20-4). Comprehensiveness means to take a liberal sample (5:20-4).

The IG must differentiate between individual aircrew members or compare an aircrew's performance to lesson objectives. There are two ways to do this. Norm referenced differentiation compares individuals to each other (5:20-4), while criterion-referenced differentiation compares each aircrew against a desired outcome or set of standards (5:20-5). The SAC IG uses criterion-referenced differentiation, comparing aircrews to the objectives of the tactics training program.

#### TYPES OF QUESTIONS

AFM 50-62 discusses two examination types. Selection tests use multiple choice, matching, and true or false questions, while supply tests use completion, short answer, or essay questions. Both types of questions are applicable to IG testing.

Selection tests offer three advantages to evaluators. They are easy to grade, statistically significant, and allow the evaluator to ask many questions in a short time period (5:21-1). Selection tests work well at low cognitive levels. The main disadvantage with selection tests is crewmembers can guess answers with a reasonable chance of being correct, especially if one or two false answers are eliminated (5:21-1).

SAC's test development guidance, SACR 50-6, builds guessing bias into tests. Though primarily a standardization regulation, SACR 50-6 is familiar to the officers at SAC Headquarters writing tactics tests. When questioned about the guidance action officers use to write tests, the responsible office at SAC Headquarters replied, "we generally follow [SACR] 60-4." SACR 60-4 is the primary standardization regulation, familiar to aircrew instructors. If an action officer uses this familiar regulation instead of AFM 50-62, it is logical to assume these officers also use SACR 50-6 to construct test questions.

SACR 50-6 states, "Each question will contain exactly four alternatives" (7:4-1). "One alternate (distractor) should represent a plausible answer only to a person with a vague knowledge of the subject" (7:4-1). "One alternative (distractor) must be entirely incorrect" (7:4-1). A crewmember with a vague knowledge of the subject can eliminate half of the possible answers and have a 50 percent chance of guessing the correct answer. Using true or false test questions does not eliminate guessing bias. Aircrews still have a 50 percent chance of correctly guessing the answer because there are only two possible choices.

One alternative is to use matching questions. Matching questions are useful in measuring understanding of closely related concepts or facts (5:21-5). The questions are easily graded, provide rapid feedback, and allow the IG to ask many questions in a limited amount of time.

Supply questions can be used to question aircrews. Essay questions are used when students are required to think reflectively or creatively (5:21-12), implying a high cognitive level. Essay questions are difficult to score because the grader needs tactical expertise and a complete answer key. Only a small number of essay questions can be asked during the examination period. The short answer question includes features of both the completion and the essay item (5:21-11), and is used to measure the ability to recall facts, basic concepts, and principles" (5:21-11). This type of question tests ideas more fully than selection questions since the recalled idea must be stated in complete form (5:21-11). It too needs a complete answer key and a knowledgeable grader. Completion items almost eliminate guessing (5:21-9), and when compared to essay questions, the evaluator can ask more questions in a given examination period.

Completion items work well at lower cognitive levels. An evaluator is able to ask many questions and grade them quickly with a simple answer key. These questions eliminate aircrew guessing by providing no possible choices, only a blank to write the answer in. Completion questions present fewer challenges to the IG than other types of supply questions.

There are challenges associated with using essay and short answer tests. The majority of answers on essay tests may not answer the question fully but demonstrate some subject knowledge. If partial credit is given the evaluator must give credit without sparking lengthy debates between the grader and the wing. Even with a knowledgeable grader and a complete grading guide, essay grades remain subjective, taking a long time to evaluate because each answer has to be read and compared to the grading guide. An evaluator can subconsciously introduce bias by evaluating writing style or penmanship instead of tactical knowledge.

#### SUMMARY

This chapter defined learning and discussed using the cognitive domain to quantify mental activities into learning objectives. Crewmembers demonstrate learning by responding to criteria-referenced examinations. The author examined characteristics of good evaluations, types of test questions described by AFM 50-62, the advantages and disadvantages associated with different test questions, and barriers placed on effective testing by staff officers using familiar regulations and instructor experience.

## Chapter Four

### HOW OTHERS EVALUATE

#### US Army

The Army tests each pilot during his annual flight evaluation. Each Army unit's evaluation pilots develop locally administered tests from a central question data base maintained by the Fort Rucker aviation school. The test consists of multiple choice questions dealing with aircraft systems, instrument flying, emergency procedures, tactics, and aircraft recognition. Passing is 80 percent.

During the ORI, units deploy into the field to demonstrate wartime capabilities. Highly qualified pilots from Fort Rucker evaluate helicopter crew knowledge through informal oral questioning dealing with any helicopter subject. There are no written tests or simulator evaluations during the ORI phase (18:--).

#### US Navy

The Navy uses multiple choice, fill in the blank or, short answer questions with all squadron personnel taking the same test at the same time. The majority of test questions deal with tactical applications, and aircraft recognition. During the inspection, the air wing is tasked to plan a strike mission graded on terrain use, force composition, and tactical considerations based on the threat. The shore based simulator is not used because the unit is evaluated at sea. Evaluators informally question flight crews during mission briefings and debriefings. Evaluator pilots can fly the wingman position to evaluate an individual's flying proficiency. The Navy evaluators are highly knowledgeable in tactics, tactical applications, and aircraft flown by the unit (16:--).

#### MAC

MAC headquarters writes a 40 question, multiple choice test, dealing with tactics, aircraft systems, and performance data. Passing is 85 percent. The simulator is not used in evaluating tactics. Evaluators informally question 40-50 percent of the aircrews flying an ORI mission during mission briefing

or debriefing asking questions from any area graded during an annual flight evaluation as well as tactics (17:--).

#### TAC, USAFE, and PACAF

TAC/AAC/USAFE/PACAF Regulation 60-2, Volume I, contains guidance for aircrew testing. This regulation establishes the command master question bank, a question data base used to generate tests (14:5-1). The TAC/USAFE/PACAF IGs test aircraft system knowledge, aircraft recognition, tactical knowledge, and technical order critical procedures. Passing score is 85 percent (14:5-3) except critical procedures that require a perfect score (14:5-3). Of 50 questions, at least half are devoted to tactics. The evaluators are augmented by evaluator pilots and intelligence officers from Numbered Air Force Headquarters. The simulator is not used in any phase of evaluation.

The majority of the evaluation occurs during the inflight phase. Aircrews are informally questioned when an evaluator flies with pilots or in their formations. The IG evaluator grades inflight threat detection and reaction from chase planes or the fighter's second seat. Tactical evaluators task other fighter units to perform aggressor/interceptor duties, attacking evaluated fighters proceeding to the bombing range. The IG uses gun cameras and voice recorders to score these interceptions and actual weapon releases to grade weapon reliability (15:--).

#### SUMMARY

Our sister services and MAJCOMs possess testing vehicles applicable to the SAC evaluation process. Each headquarters, except MAC, tests aircraft recognition. No one uses the aircraft simulator and tests contain many selection questions. The Navy and MAC use mission planning exercises to measure the ability to apply tactical knowledge. All evaluators informally question the aircrew they fly the ORI mission with. The evaluators are highly competent in the material they will evaluate. The tactical forces emphasize inflight mission performance. SAC can use some of these testing vehicles to improve the command's IG testing program.



## Chapter Five

### NON-WRITTEN EVALUATION

#### FORMAL ORAL EVALUATIONS

The mission certification is SAC's formal oral evaluation. The aircrew briefs a specific bomber mission to a panel comprised of a colonel and representatives of the wing's mission support agencies. The briefing lasts for approximately two hours, with one hour for the aircrew briefing and an additional hour for the staff officers to question the aircrew to cover omitted information and re-emphasize important points. The certification forces aircrews to operate at higher cognitive levels by applying their basic knowledge to sortle specific challenges. The certification eliminates guessing while allowing the IG to determine if aircrews are meeting the objectives of the tactics program specified in SACR 3-1, Volume II.

The IG faces challenges in administering a certification program. The IG must anticipate the wing volunteering only the best aircrews to brief, forcing the IG to randomly select certifying aircrews to insure an aircrew cross section is evaluated. Bringing together the large number of people required for a certification may create physical space problems. At a certification, in addition to the IG representatives, the wing staff will want to be represented. There are very few work areas on a base meeting the security criteria required for the briefing. In the author's experience, only the briefing room at the war plans center will be available. Requiring only one briefing assumes a small sample size to be representative of the wing while eliminating most space constraints. Requiring more than one briefing gives a better indication of aircrew ability but may turn the certification process into an all day affair because of space and IG personnel limitations. The IG may not have the time or evaluators available to conduct certifications. In the last two ORIs the author witnessed, no mission certifications were administered.

This author believes certifications are an important part of the overall evaluation process. It allows the aircrews to translate book knowledge into concepts and relate those concepts to real world challenges. Aircrews are challenged to provide an oral essay answer to a panel of experts, eliminating the time and grade keys required for a written essay test. The certification program is already in existence, and needs only to be employed during every evaluation. The challenges in implementing this already in-place program are administrative and could be easily overcome by prior planning.

## MISSION PLANNING

The means to evaluate aircrew mission planning activities exists in SAC's IG system. Tasking a bomber crew to replan a sortie is done on a selective basis, depending on a wing's taskings. Bomber crews must be capable of plotting target positions, air refueling routes, and ingress and egress routes, taking into account the tactical environment. The aircrew assembles the administrative paperwork to support their mission planning and submits the entire package to the IG for evaluation.

The aircrew is operating at a high cognitive level by applying knowledge level concepts to the problem. This mission planning exercise could be administered to each wing. A tactically qualified individual will evaluate the aircrew's product along with the agencies currently tasked to evaluate the plan.

## USING THE SIMULATOR

Using the weapons system trainer (WST) seems to provide an excellent opportunity for the IG to observe aircrews perform maneuvers and react to threats in a manner precluded by peacetime flying safety. The aircrews are performing at a high cognitive level. The IG could develop a selection of maneuvers to evaluate each aircrew, quantify aircrew performance into a grade, and issue a report.

This author believes the WST does not have a place in IG evaluations of aircrew tactical knowledge. SACR 51-52 states the purpose of the WST is "to practice tactics and malfunction/emergency procedures". It does not address using WSTs for evaluation (8:7-1). Every base does not have a WST forcing some wings to commute aircrews to a WST site during an intense period of evaluation. IG evaluators may not be current or qualified in the B-52, or have little or no evaluation experience. The author believes the best choice for evaluators would be the command level evaluators, 1st Combat Evaluation Group (CEVG). The members of CEVG possess the evaluative skills, are the best aircrews SAC has, and are current in B-52 operations.

CEVG and the WST are already involved in evaluating the tactics program. SACR 60-4, Volume I, tasks CEVG to standardize SAC's tactics (11:11-1). CEVG visits each wing approximately every 24 months (11:11-1) to provide instruction to the wing's aircrews (11:11-1) using the WST (11:11-1). Though not a formal inspection (11:11-1), CEVG debriefs squadron commanders about unit strengths and weaknesses. Because the WST is already used in the overall evaluation of the tactics program, the author believes it is redundant to evaluate a wing's aircrews in the WST during the ORI.

## INFORMAL ORAL EVALUATIONS

Other commands use informal oral evaluations extensively in the flying evaluation phase. Informal oral evaluations allow aircrews to operate at

higher cognitive levels. The crewmember being questioned cannot guess. The IG does not need to gather a large group of evaluators together as in the formal briefing. The rating associated with this questioning is hard to quantify and subject to evaluator bias.

If the IG accepts formal oral evaluation results as evidence of aircrews meeting training objectives, then requiring informal oral evaluations only re-evaluates aircrew ability to apply basic tactical knowledge. Unlike a formal oral evaluation, there is no mechanism in place to conduct this type of interview, and a single evaluator instead of a panel, introduces possible evaluator bias. Because of the subjective nature of the rating, possible questioning of evaluator judgement in awarding ratings, the need for complete grading guides, and in place formal oral evaluations, the author believes informal oral evaluations should be limited to questions clarifying inflight events and data recording by the IG.

### INFLIGHT EVALUATION

The aircrew flying a B-52 is operating at the evaluation cognitive level, "exercis[ing] learned judgement" (5:2-2). There is no guess work as the crew translates what they know into action. Flying the B-52 is the ultimate test of an aircrew's capability to perform a mission simulating their Emergency War Order sortie. The author feels SAC can place greater emphasis on inflight evaluation.

Not enough evaluators fly with the aircrews during ORI missions. In the author's experience, usually two members of the IG team fly with aircrews during the ORI flying phase. These evaluators fill out data sheets used in computing the DE grade. Using two evaluators does not provide a comprehensive examination of the wing's aircrew and aircraft. Some discrepancies applicable to calculating the DE rating do not get reported. Low numbers of evaluators allow aircrews and wings to "game" the DE system by not powering certain aircraft systems that lose points upon failing or by not writing malfunctions up in the aircraft records that are checked by the IG for information to calculate damage expectancy.

This author believes enough knowledgeable individuals should be assigned to the IG team to increase the inflight evaluation rate to 50 percent of the sorties a wing flies. This will provide a comprehensive look at aircrew ability, discourage gaming the evaluation system, provide accurate feedback on systems capabilities, and provide an uninflated DE rating.

SAC's IG team should increase its manning rather than augmenting the IG team with officers from the headquarters tactics staff. Using tactics officers as augmentees detracts from their primary job of managing SAC's tactics initiatives. Adopting the author's recommendation will require creation of additional manpower allocations allowing headquarters staff officers to fly.

There are more activities the IG can evaluate inflight to measure

the ability to flight. Actually releasing inert weapons is being studied by the SAC staff and will add a great deal of realism to evaluations. Decreasing allowed communications, flying at night as much as possible, and grading weapon release parameters are three areas that could be integrated immediately into the current evaluation system with little difficulty.

#### SUMMARY

The author believes formal oral evaluation, mission planning activities, increasing the number of sorties carrying IG evaluators, and challenging aircrews through increased inflight evaluation will allow SAC to better evaluate aircrew abilities. Aircrews operate at high cognitive levels, guessing is eliminated, and the means to implement these programs is already in place in the evaluation system. The author feels informal oral evaluation and using the simulator should not be used because they are redundant, are not in place in the current evaluation system, and could be subject to evaluator bias because of a single evaluator.

## Chapter Six

### SOLUTIONS AND RECOMMENDATIONS

#### TESTS

The author believes SAC must test tactics separately from EWO knowledge. Tactics and EWO knowledge are not the same thing although historically SAC treated tactics as a segment of EWO knowledge. The changes to the testing program brought by the publication of SACR 3-1 did not separate or increase the number of tactics questions. Those questions were merely shifted to a different examination phase. A separate tactics test allows evaluators to incorporate the effective test characteristics the author discussed in Chapter Three.

Separate tests measure exactly what SAC desires to measure, either tactical knowledge or EWO knowledge. A separate tactics examination of 25 questions allows for a more comprehensive sample than the six tactics questions in today's examinations (12:10-74). The six tactics questions in the EWO knowledge test could be converted to EWO knowledge questions, improving the comprehensiveness of that examination without lengthening it.

A new grading area, "tactical knowledge", needs to be created to reflect the results of the separate testing program, emphasize the importance of tactics, and provide an area where other tactical grading results could be summarized. The author sees tactical knowledge as a grading area of equal stature with the current "aircrew knowledge" rating.

A technically correct test, written by SAC's academic experts, using the criteria in AFM 50-62, will provide the best test results. Test questions must be referenced to tactics lesson's learning criteria. Multiple choice, matching, or completion questions should be used because they relate cognitively to the knowledge level objectives of the tactics program as well as allowing evaluators to ask many questions in a short time.

SAC should create a command master question bank for tactics questions. This data base can be updated insuring current test questions reflecting the latest Soviet capabilities. Once created, tests do not repeatedly have to be written. Coupled with a random question generating program, the IG can create any number of different tests reducing the probability of a single test compromise.

The author feels the 436th Statagic Training Squadron (STS) is best suited to create the master question bank. As the authors of all SAC ground training lessons, the 436th STS is familiar with the objectives of each lesson in the tactics program. As the 436th STS develops lessons it would write test items, enter the questions into a data base, review the data base for technical accuracy, and forward it to the IG. Having professional lesson developers write tactics tests releases SAC tactics action officers from this additional duty allowing them to concentrate on their primary job of staffing tactics initiatives.

The ability to tell friend from foe is important in combat. Testing aircraft recognition further challenges aircrews while insuring they can recognize threats. Five 35mm slides could be shown to aircrews who must identify the threat using NATO code-names. These slides should provide an aircraft picture easy to identify without being obvious. Testing everyone encourages each crewmember to be familiar with the threat, threat capabilities, and bomber defensive actions.

#### NON-WRITTEN EVALUATION

The author recommends certifying two aircrews in each wing. Two crews provide a more comprehensive sample but must take into account the facilities limitations discussed in Chapter Five. The certifications would be worth 20 percent of the new tactical knowledge grading area.

The author further recommends the IG choose an aircrew to replan a bomber sortie during each inspection. After the plan is complete it would be evaluated using both tactical and existing criteria. The rating worth 10 percent of the new tactical knowledge grading area would be awarded. A crew certifying a mission would not be allowed to conduct the mission planning, enabling the IG to evaluate more aircrews.

#### INFLIGHT EVALUATION

Increasing the number of inflight evaluators contributes to a comprehensive evaluation of random factors measured by the DE grading criteria. There would be increased evaluation of altitude maintenance, mission timing, and route adherence. Equipment serviceability, which weighs heavily in DE computations, can be more accurately measured using IG reports and not wing aircrew or maintenance reports. This would provide a more complete profile of crew force capabilities.

The Soviet's radio detection capability is formidable. In order to increase our tactical advantage, the author recommends flying ORIs using increased communications security (COMSEC) and emission control (EMCON) procedures. SACRs 3-1 and 51-52 provides guidance for refueling and formation flying in a simulated EMCON environment.

Today the only ORI communication restriction is "crews will maintain

radio silence after the HHCL\* (17:10-19). The training event defined in SACR 51-52 by the computer identifier R-057 can be used during the rendezvous and refueling. R-057 provides for one radio contact 15 minutes prior to the air refueling control time, and refueling under radio silence using visual signals (8:A2-15-A2-16). The B-52 training event defined in SACR 51-52 by the computer identifier P-109 simulates the restricted communication environment while flying bombers in formation. The bomber formation must fly using emission restrictions specified in SACR 3-1 for the entire mission including the preflight (8:A2-12). Using these two training events effectively provides a realistic flying scenario by simulating a wartime EMCON environment.

The author recommends flying as many ORI missions at night as possible to challenge the aircrews in an adverse environment. SAC doesn't fly a great deal at night. The author's flying logbook shows less than 25 percent of his 2000+ hours in the B-52 is nighttime. Many young aircraft commanders have less night experience. Air refueling at night is a challenge, even for an experienced aircraft commander. In the author's experience as a chief instructor pilot, night air refueling is a potential weak area for young B-52 aircraft commanders. Visual bombing and navigation references are lost at night, simulating a cockpit with raised thermal protection curtains and fatigue becomes a factor. These challenges more realistically simulate a wartime environment as well as getting young aircrews the night flying practice they need.

The author recommends IG inflight evaluators record weapon release parameters for post flight comparison to technical order limits. This was suggested by 15th Air Force's Tactics Office (20:--). Weapons are delicate devices. If the pilot does not attain release parameters, the bomb will skew, collide with the aircraft, or possibly be a dud. Aircraft commanders are usually trying to center the steering indicator (FCI) right up to weapon release. While the bomb scoring site doesn't know if your aircraft was in a bank when computing the bomb score, the weapon would know. Recording bank, pitch, and altitude at weapon release will allow better assessment of probable damage from that weapon, producing a better estimate of actual damage expectancy.

### RECOMMENDATIONS

The author recommends that the Strategic Air Command implement the recommendations made in this paper while continuing to explore new areas to be evaluated inflight to produce a realistic combat environment. If expected fiscal or manpower reductions materialize, the author recommends implementing those recommendations not affected by the reductions and implementing the remainder of the recommendations as soon as funding becomes available.

# BIBLIOGRAPHY

## Books

1. Denova, Charles C. TEST CONSTRUCTION FOR TRAINING EVALUATION. New York, NY: Van Nostrand Reinhold Company, 1979.
2. Taylor, John W.R. ed. JANE'S ALL THE WORLD'S AIRCRAFT. New York, NY: Jane's Publishing Co, 1985.
3. ----- WEBSTER'S NEW COLLEGIATE DICTIONARY. Springfield, MA: G & C Merriam Co, 1974.

## OFFICIAL DOCUMENTS

4. ----- JCS Publication 1, DEPARTMENT OF DEFENSE DICTIONARY OF MILITARY AND ASSOCIATED TERMS. Washington D.C., U.S. Government Printing Office, 1 June 1987.
5. United States Air Force. AFM 50-62, HANDBOOK FOR AIR FORCE INSTRUCTORS. Randolph AFB TX: ATC/OC, 15 January 1984.
6. Strategic Air Command. SACR 3-1, Vol V., B-52 TACTICS. Offutt AFB NE: SAC/DOSJ, 27 July 1987. SECRET. Classified by multiple sources. Declassify: OADR.
7. Strategic Air Command. SACR 50-6, STANDARDIZATION AND EVALUATION CONCEPTS AND TECHNIQUES. Offutt AFB NE: SAC/DOITA,
8. Strategic Air Command. SACR 51-52, B-52 AIRCREW TRAINING. Offutt AFB NE: SAC/DOITA 28 September 1987. 4 December 1978.
9. Strategic Air Command. SACR 55-21, Vol. II, B-52 AIRCREW TACTICAL DOCTRINE. Offutt AFB NE: SAC/DOJT 18 June 1986. SECRET. Classified by multiple sources. Declassify on OADR.
10. Strategic Air Command. SACR 55-45, Vol. V, AIRCREW EMERGENCY ACTION PROCEDURES. Offutt AFB NE: SAC/DOCC, 16 June 1987. SECRET. Classified by multiple sources. Declassify on OADR.
11. Strategic Air Command. SACR 60-4, Vol. I, STANDARDIZATION/EVALUATION PROGRAM. Barksdale AFB LA: 1 CEVG/DOV, 19 Oct 1987.
12. Strategic Air Command. SAC SUP 3 to AFR 123-1, THE INSPECTION SYSTEM. Offutt AFB NE: SAC/DOIN 20 June 1986.



## CONTINUED

13. Strategic Air Command. SACRP 127-1, COMBAT CREW. Vol. XXXV, No.6  
Offutt AFB NE: SAC/IGFE
14. Tactical Air Command. TAC/AAC/USAFE/PACAFR 60-2, Volume 1,  
AIRCREW STANDARDIZATION/EVALUATION PROGRAM. Langley AFB, VA:  
TAC/DOV 24 April 1987.

### OTHER SOURCES

15. Eberhart, Jeff, Major, USAF. ACSC Student. TAC's evaluation  
system. Interview.
16. Hillier, David, Lt Cmdr, USN. ACSC Student. USN's evaluation  
system. Interview.
17. Olsen, Dean, Major, USAF. ACSC Student. MAC's evaluation system.  
Interview.
18. Travis, Kenneth, Major, USA. ACSC Student. US Army's evaluation  
system. Interview.
19. Hanfland, Larry. Lt. Col, USAF. SAC Background Paper "SACR 3-1" IG  
testing". 1 June 1987
20. Oldham, Huel, Col, USAF. 15th AF tactics initiative "evasive  
action bombing" 21 August 1987
21. Welch, Larry, General, USAF. Memo to Major General Beckel, SAC/DO.  
Subject: poor state of SAC tactics program.

END  
DATED  
FILM  
8-88  
Dtic